SUPPLEMENTARY MATERIALS

Author Contributions

Materials and Methods

Figures S1-S10

Tables S1-S7

References (*1-50*)

AUTHOR CONTRIBUTIONS

JT, LBB, MEW, and ZPJ designed the study. NSM, JNK, AOS, SM, and JFB collected the data. NSM, JS, JT, LBB, JCG, and RPR analyzed the data. JT, LBB, NSM, and JS wrote the paper, with contributions and edits from all authors. JT, LBB, MEW, ZPJ, and NSM provided funding support.

MATERIALS AND METHODS

Text S1. Study subjects and experimental design

Study subjects were 45 adult female rhesus macaques housed in groups of five females each at the Yerkes National Primate Research Center Field Station (Table S1). Nine groups were initially formed in January – June 2013 (Phase 1), with group membership rearranged to form 9 new groups in March – June 2014 (Phase 2). Group construction followed a previously established protocol (*12*) in which females were sequentially introduced into indoor-outdoor run housing (25 m by 25 m for each area) over the course of 2 - 15 weeks. In both phases, behavioral data collection for each group started after the fifth and final female was introduced into the group; biological sample collection started 10-16 weeks thereafter (median = 12 weeks) to ensure that rank hierarchies were completely stabilized. To maximize between-phase changes in dominance rank, each social group in Phase 2 consisted of females who shared the same dominance rank or (due to the uneven number of social groups in our study) were separated by a maximum of 1 ordinal rank difference in Phase 1 (Table S1). In all cases, no co-housed females were close kin (e.g., parent-offspring, grandparent-grandoffspring, full sibling or half sibling: fig. S8), and the vast majority (97%, 174 of 180 dyads) had never seen each other prior to being co-housed in this study (in either phase).

To monitor dominance rank, we collected behavioral data using weekly focal observations (345 total hours of observation, 223.5 hours in Phase 1 and 121.5 hours in Phase 2). Except where noted, we assigned dominance rank using Elo ratings, a continuous measure of rank in which higher scores correspond to higher status and ratings are updated following any interaction in which a winner or loser can be scored (*13, 24, 25*). Wins increase the winner's Elo rating and losses decrease the loser's Elo rating, proportional to the expected outcome of the interaction. Thus, when a higher ranking female wins an encounter against a low ranking female, her Elo rating changes little because the result was expected based on her pre-interaction rating relative to her opponent's. However, if the outcome were reversed, both females' Elo ratings would change substantially to reflect the observed rank dynamics.

We calculated Elo ratings using all dominance interactions that occurred after the fifth and final female was introduced into a social group. For all females, the initial Elo rating was set to a value of 1000 and the baseline number of points a female could potentially gain or lose during a dominance interaction (k) was set to 100. For each interaction, k was weighted by the expected probability of an individual winning or losing, based on a logistic function that was updated after each dominance interaction (*24*). In both phases, order of introduction predicted subsequent dominance rank (Phase 1: $r = -0.57$, $p = 4.1 \times 10^{-5}$, $n = 45$ females; Phase 2: $r = -0.68$, $p = 3.3 \times 10^{-7}$, $n = 45$ females, based on values at the time of sampling for cell type-specific expression analyses; Fig. 1B), and Elo ratings remained highly stable for the duration of each phase (Elo stability index range: 0.995 – 1.00; where 1 represents a perfectly stable hierarchy in which lower-

ranking females never outcompete higher-ranking females (*26*)). Elo ratings were also highly correlated with more traditional ordinal measures of dominance rank (Pearson's r = -0.94, $p = 6.1x10^{-40}$ (*17*)).

To test whether Elo ratings predicted social interactions, as expected, we assessed rates of received harassment and grooming. Received harassment was defined as the per hour rate of harassment received from a female's group mates, mean-centered across groups so that the average rate of harassment received by females in a group was set equal to 0. Grooming rates were defined as the amount of time per hour a female spent grooming with her group mates, also mean-centered to 0 for each social group. We also used these values to test the mediating role of agonistic (received harassment) and affiliative (grooming) behaviors on rank effects on gene expression levels, for the cell type-specific gene expression analysis (see "Behavioral mediation analysis" below).

Text S2. Blood sample collection

To measure gene expression levels in purified cell populations, we drew 12-20 mL of blood from each female, purified the PBMC fraction using density gradient centrifugation, and performed fluorescent activated cell sorting on a BD FACSAria IIu machine housed at the Duke Human Vaccine Institute Flow Cytometry Core. We sorted cell types as follows: classical monocytes ($CD3^-/CD20^-/HLA\text{-}DR^+/CD14^+$), natural killer cells ($CD3^-/CD20^-/HLA\text{-}DR^-/CD16^+$), B cells ($CD3^-/CD20^+/HLA\text{-}DR^+$), helper T cells ($CD3^+/CD8^-/CD4^+$), cytotoxic T cells ($CD3^+/CD8^+/CD4^-$) (see fig. S3 for sorting strategy and Table S2 for antibody information). Purified cells were immediately pelleted, lysed in buffer RLT, and frozen at -80°C until DNA and RNA extraction using the Qiagen AllPrep extraction kit. Sorted cell populations (five populations per blood sample) were obtained once for each study subject in each of the two study phases.

To test how dominance rank affects the LPS response, we drew 1 mL of whole blood from each female (in Phase 2 only) directly into a TruCulture tube (Myriad RBM) that contained cell culture media plus 1 μg/mL ultra-pure lipopolysaccharide (LPS) from the *E. coli* 0111:B4 strain ("LPS+"), and another 1 mL of blood from the same draw into a TruCulture tube that contained cell culture media only ("control"). Samples were incubated in parallel for 4 hours at 37°C. We then separated the serum and cellular fractions, and lysed and discarded the red cells from the remaining cell pellet by applying red blood cell lysis buffer (RBC lysis solution, 5 Prime Inc.) for 10 minutes followed by centrifugation. We lysed the remaining white blood cells in Qiazol for storage at -80°C, and extracted total RNA from each sample using the Qiagen miRNAEASY kit. To control for variation in cellular composition in downstream analyses, we also used flow cytometry to quantify the proportion of 11 different cell types in blood samples obtained in the same draw: polymorphonuclear ($CD14^{dim}$/SSC-A>100K/FSC-A>100K), classical monocytes ($CD14^+/CD16^-$), $CD14^+$ intermediate monocytes ($CD14^+/CD16^+$), $CD14^-$ non-classical monocytes ($CD14^-/CD16^+$), helper T cells ($CD3^+/CD4^+$), cytotoxic T cells ($CD3^+/CD8^+$), double positive T cells ($CD3^+/CD4^+/CD8^+$), $CD8^-$ B cells ($CD3^-/CD20^+/CD8^-$), $CD8^+$ B cells ($CD3^-/CD20^+/CD8^+$), natural killer T lymphocytes ($CD3^+/CD16^+$), and natural killer cells ($CD3^-/CD16^+$) (fig. S9).

Text S3. RNA-seq data generation

Library construction. We constructed RNA-sequencing libraries for each sample

(purified cell populations in both phases and control and LPS+ samples in phase 2) using the NEBNext Ultra RNA Library Prep Kit (New England Biolabs). Briefly, we purified the poly-adenylated mRNA from 200 ng of total RNA using the NEBNext Poly(A) mRNA Magnetic Isolation Module. The mRNA was then reverse transcribed into cDNA, ligated to Illumina adapters, size-selected for a median size of ~350 bp, and amplified via PCR for 13 cycles. We tagged each sample with a unique molecular barcode and pooled 10-12 samples per Illumina HiSeq 2500 lane of single-end 100 bp sequencing.

Read alignment. Following sequencing, we trimmed Illumina adapter sequence from the ends of the reads using Trim Galore! (v0.2.7), mapped trimmed reads to the rhesus macaque genome (*MacaM* v7; (*27*)) using the STAR 2-pass method (*28*), and collated the number of reads that mapped uniquely to each annotated *MacaM* gene (v7.6.8) using HTSeq-count (v0.6.1) with the option "intersection-nonempty" (*29*). For the LPS challenge experiment and each cell type-specific data set, this procedure resulted in a p x n read-count matrix, where p is the number of genes measured and n is the number of samples in the given data set. After excluding samples that did not produce sequenceable libraries and post-sequencing quality control, we generated read counts for 83 samples in the LPS challenge experiment (43 controls and 40 treatment), and 440 samples for the cell type-specific analysis (89 in helper T cells, 89 in cytotoxic T cells, 86 in B cells, 88 in natural killer cells, and 88 in monocytes). We confirmed the identity of all samples based on genotyping from the RNA-seq reads (see below for details).

Text S4. Read normalization and correction for batch effects

Prior to RNA-seq data analysis, we first filtered out genes that were very lowly or not detectably expressed in our samples. For the cell type-specific data sets, we removed any gene with a median RPKM ≤ 2 in that cell type. For the LPS challenge experiment, genes were removed if the median RPKM was ≤ 2 in both control and LPS-stimulated samples. This procedure resulted in a different number of genes that passed filter in each data set: 8,823 genes in helper T cells, 8,802 genes in cytotoxic T cells, 8,388 genes in monocytes, 8,754 genes in natural killer cells, 8,951 genes in B cells and 9,047 genes for the LPS experiment (control and LPS+).

For each of the six data sets, we normalized the resulting read count matrix using the function *voom* from the R package *limma* (*30*). We then modeled the normalized expression values as a function of the sample donor's social group membership (9 social groups in each phase for the cell type-specific data sets, for a total of 18 distinct groups; and the 9 social groups in Phase 2 only for the LPS challenge data set). Because we sampled all females in each group within a short time frame (median time frame for the complete group to be sampled was 3 days; groups were entirely sampled for a given analysis within a 10 day window, with all but two exceptions when the final female was sampled several weeks later), controlling for the identity of the group removes biological variation related to differences in group dynamics, and most technical batch effects related to sample collection and processing. Thus, we used the residuals of the model relating normalized expression to social group identity as our primary outcome measure in all subsequent analyses.

Text S5. Genotyping

We used genotype data to confirm sample identity across experiments and to

control for genetic relatedness among individuals in our analyses (pairwise genetic relatedness in our sample was on average low, but several close kin were included in the data set, although never housed in the same social group: fig. S8) To do so, we combined the RNA-seq reads for all five purified cell types for each female and called variants using HaplotypeCaller from the Genome Analysis Toolkit (GATK v3.3.0), following the Best Practices for variant calling in RNAseq (https://www.broadinstitute.org/gatk/guide/article?id=3891). After genotyping we retained sites that passed the following filters: quality score ≥100; QD < 2.0; MQ < 35.0; FS > 60.0; HaplotypeScore >13.0; MQRankSum < −12.5; and ReadPosRankSum < −8.0. Finally, we estimated kinship with the program lcMLkin (*31*) using single nucleotide variants that were genotyped in all 45 females and that were thinned to be at least 10 kb apart (n = 54,165).

Text S6. Modeling rank effects on gene expression

To identify genes that were significantly affected by dominance rank, we used a linear mixed effects model that controls for relatedness within the sample (*32-34*). We analyzed each of the six data sets (five purified cell types plus the LPS challenge data) separately using the R package EMMREML (*35*), in each case working with gene expression values after controlling for social group/batch effects, as described above.

For each gene in the cell type-specific data sets, we first estimated the effect of dominance rank on gene expression levels across phases using the following model:

$$\begin{gathered} y = \mu + r\beta + a\gamma + Zu + \varepsilon, \\ u \sim MVN(0, \sigma_u^2 K), \\ \varepsilon \sim MVN(0, \sigma_e^2 I) \end{gathered} \tag{1}$$

where y is the n by 1 vector of residual gene expression levels for the n samples collected in Phase 1 and Phase 2; μ is the intercept; r is an n by 1 vector of Elo ratings and β is its effect size; and a is an n by 1 vector of female age in years at the time of sample collection and γ is its effect size. The m by 1 vector u is a random effects term to control for kinship and other sources of genetic structure. Here, m is the number of unique females in the analysis (m=45), the m by m matrix K contains estimates of pairwise relatedness derived from a 45 x 54,165 genotype data set, σ_u^2 is the genetic variance component (0 for a non-heritable trait, but most gene expression levels are heritable: (*36, 37*)), and Z is an incidence matrix of 1's and 0's that maps measurements in Phase 1 and Phase 2 to individuals in the random effects term (thus controlling for repeated measurements for the same individual across phases). Residual errors are represented by ε, an n by 1 vector, where σ_e^2 represents the environmental variance component (unstructured by genetic relatedness), I is the identity matrix, and MVN denotes the multivariate normal distribution. For each data set and gene, we tested the null hypothesis that $\beta = 0$ versus the alternative hypothesis, $\beta \neq 0$. To test for dominance rank effects on the cellular composition of our samples (based on flow cytometry data), we used a parallel model replacing gene expression level (y) with the logit-transformed proportional representation of each of 11 cell types and added a fixed effect term to account for three rounds of repeated measures for the flow data (once in Phase 1 and twice in Phase 2; fig. S2). As in eq (1), this model included a random effects term to control for repeated

measurements from the same genotype/individual.

To test for consistency and plasticity of rank effects on gene expression levels *across* phases, we also ran a nested model that estimated the effects of rank specific to each phase:

$$y = \mu + r_p\beta_1 \times I(p = 1) + r_p\beta_2 \times I(p = 2) + a\gamma + Zu + \varepsilon, \quad (2)$$

where the notation is the same as above, but I is an indicator variable for phase, p (phase 1 or phase 2). To produce an empirical null distribution for rank effects in these two models, we performed 1000 permutations of female dominance rank and re-ran the analyses using randomized rank values.

To test for rank effects on gene expression in the LPS challenge experiment, we estimated the effects of rank specific to the control and treatment conditions (all data were collected for Phase 2 only):

$$\begin{aligned} y = \mu + \delta c + r\beta_1 \times I(c = 0) + r\beta_2 \times I(c = 1) + XV_1 \times I(c = 0) \\ + XV_2 \times I(c = 1) + Zu + \varepsilon \end{aligned} \quad (3)$$

where the notation for gene expression, dominance rank, the intercept, residuals, and random effects are consistent with the models above, but I is an indicator variable for c, the condition (control = 0, treatment = 1), rather than for phase, and an additional effect, δ, is fit for the binary condition variable. Other fixed effects covariates, for mean-centered age and the first two principal components from a PCA decomposition of the normalized cell type proportion data (matrix X), are also estimated as nested effects within condition (vectors V_1 and V_2). To explicitly extract interaction effects between dominance rank and condition, we reformulated model (3) as follows:

$$y = \mu + \delta c + r\beta + XV + (r \times c)\beta_{rxc} + (X \times c)V_{X \times c} + Zu + \varepsilon, \quad (4)$$

where we tested the null hypothesis that $\beta_{rxc} = 0$ versus the alternative hypothesis, $\beta_{rxc} \neq 0$. For the LPS challenge experiment, we produced empirical null distributions for rank effects by permuting the dominance rank values associated with pairs of samples (control and LPS-treated) as a block, 1000 times, and re-running our analyses. To produce a corresponding empirical null for the effect of LPS stimulation (i.e., condition) on gene expression, we randomized condition within each pair of samples for each female. When only a treatment or control sample was available for a female, we randomly assigned it to one of the two conditions with 50% probability.

For all of our analyses, we controlled for multiple testing using a generalization of the false discovery rate method of Storey and Tibshirani (*38*) to empirical null p-value distributions generated via permutation (code available upon request).

Text S7. Gene Ontology enrichment

Gene ontology (GO) enrichment analyses were performed using the Cytoscape module ClueGO (*39*). We conducted one-sided Fisher's Exact Tests for enrichment and corrected for multiple tests using the Benjamini-Hochberg (B-H) method (*40*). To reduce the multiple testing space and account for the nested nature of GO terms, we analyzed

only terms that fell between levels 3 and 8 of the GO tree in the term category of Biological Process, included at least 10 genes in our data set, and for which at least 5% of the total number of genes belonging to the GO term were present in the test gene set. We report significant terms as those that were enriched in the target gene set at a B-H corrected p-value ≤ 0.05 (Table S4). In the network diagrams, this threshold was sometimes made more stringent for the sake of visualization (i.e., in Fig. 3D we only show terms enriched in category I at FDR $< 10^{-5}$, while, in fig. S4A, we restrict the results shown to those with a FDR < 0.01).

For the five cell type-specific data sets, we conducted enrichment analyses separately for genes that were more highly expressed in low-ranking animals and genes that were more highly expressed in high-ranking animals, in both cases against the background set of all measured genes in that cell type. For the LPS challenge data, we conducted one enrichment analysis for each of the four LPS-rank interaction categories (stratified by direction of the LPS response and direction of the rank effect on gene expression levels in the LPS+ condition, as reported in the main text; Fig. 3D) against the background set of all genes that responded significantly to LPS at an FDR < 0.05.

Text S8. Meta-analysis for tissue specificity

True shared effects of rank across cell types could be missed in our above analyses if they failed to survive FDR correction in some cell types. We therefore complemented our independent analyses of purified cell populations with a meta-analytic approach, MeSH (Meta-analysis with Subgroup Heterogeneity; (*41*)), focusing on the union set of 1,622 genes that were detectably expressed in all five cell types and significantly associated with rank, in any cell type, at an FDR<10%. We ran MeSH for each gene using the Exchangeable Effects (EE) model, using β, the effect of rank on a gene's expression; se(β), the standard error on that coefficient; and a grid file of discrete priors for the mean effect of rank across cell types (ω) and the heterogeneity of effects across cell types (ψ). We constructed the grid file to span the effect sizes observed in our data (ω = [0.00,0.01,0.02,0.03,0.04,0.05,0.06,0.07,0.08] and ψ = [0.000,0.002,0.004,0.006,0.008,0.010,0.012,0.014,0.016,0.018,0.020]).

Text S9. Behavioral mediation analysis

To estimate the contribution of agonistic and affiliative behaviors to the effect of social status on gene expression levels, we conducted two mediation analyses: one investigating the mediating effects of received harassment, and the second testing the mediating effects of grooming rates. We investigated only genes for which we detected a significant rank-gene expression relationship in the cell type-specific analyses, focusing on the two cell types (natural killer and helper T cells) for which this signal was strongest.

For each gene, we were interested in estimating the indirect effect of dominance rank on gene expression levels through the mediating variable (grooming or received harassment). The strength of the indirect effect was estimated as the difference between the effect of rank in two models: the "unadjusted" model that did not account for the mediator (equivalent to β in equation 1 above), and the effect of rank in an "adjusted" model (β' in equation 5 below) that incorporated the mediator, *t*.

$$y = \mu + r\beta' + t\gamma_t + a\gamma + Zu + \varepsilon, \quad (5)$$

where notations are the same as in equation (1), but with the addition of an effect for the mediating variable, γ_t. To assess the significance of the indirect effect, $\beta - \beta'$, we performed 1000 iterations of bootstrap resampling to calculate 95% confidence intervals for each mediator. We deemed an indirect effect significant if the lower bound of the 95% confidence interval did not overlap with 0.

Text S10. Transcription factor binding site enrichment

To investigate transcription factors involved in rank effects on gene expression, we combined data on the location of predicted transcription factor binding sites in the rhesus macaque genome with data on open chromatin in PBMCs collected from our study population. To identify these regions, we generated ATAC-seq libraries from three mid-ranking females sampled during Phase 2 (using 50,000 cells per female and following (*42*), but skipping the cell lysis step). Libraries were barcoded, multiplexed, and sequenced on an Illumina NextSeq 500 using paired-end, 38 bp reads. Reads were mapped to *MacaM* using the *bwa-mem* algorithm with default settings (*43*), and exhibited the typical periodic pattern of insert sizes associated with nucleosomes. Open chromatin regions were identified using MACS2 (*44*) (--nomodel --keep-dup all -q 0.05 -f BAMPE) after merging data from all three libraries into a single alignment file, at a 5% FDR threshold.

To identify putative TFBSs associated with open chromatin and rank-responsive genes, we scanned the rhesus macaque genome for matches to 1900 position weight matrices (PWM) obtained from TRANSFAC and JASPAR (*45, 46*), using a PWM model as in (*47*) and a threshold of 13 (i.e., a 2^{13}-fold increase in the likelihood that a motif is a binding site relative to background). Because multiple PWMs are often associated with the same transcription factor (TF), we clustered them by calculating the Jaccard distance between genomic locations associated with each pair of PWMs with the bedtools function *jaccard* (*48*). After hierarchical clustering of the resulting dissimilarity matrix, we considered all PWMs with a dissimilarity ≤ 0.2 (i.e., a Jaccard statistic ≥ 0.8) as members of the same "TF cluster", producing 913 non-redundant TF clusters. We also filtered out all TF clusters that were rarely found in open chromatin near annotated genes (<100 cases genome-wide), resulting in a final set of 460 motif clusters. For each gene in our data set, we then noted whether each TF cluster fell within 5 kb upstream or downstream of the gene transcription start site.

Using this information, we performed two analyses. First, we used Fisher's Exact Tests to ask whether rank-responsive genes were more likely to be associated with a given TF cluster than genes with no evidence for rank effects ($p > 0.3$). We tested for overrepresentation of TF clusters separately for genes that were more highly expressed in high-ranking females versus those that were more lowly expressed, and performed this analysis separately for the natural killer cell data set, the helper T cell data set, and the LPS challenge data set. To control for multiple hypothesis testing, we used the false discovery rate approach of Benjamini and Hochberg (*40*). Second, we used an elastic net logistic regression approach to ask about the predictive power of TFBS locations in open chromatin regions to discriminate between (i) genes upregulated in high ranking individuals versus genes unaffected by rank; (ii) genes downregulated in high ranking

individuals versus genes unaffected by rank; and (iii) genes that were upregulated versus downregulated with higher rank. To avoid biases caused by unbalanced outcome classes, we randomly subsampled genes from the majority class (without replacement) equal to the number of genes in the minority class, calculated prediction accuracy (1 – the model misclassification rate) using the function "cv.glmnet" in the R package *glmnet* (*49*), and repeated this procedure 1000 times In each iteration, we extracted the betas for each TFBS from the model with the best prediction accuracy (because the elastic net approach induces sparsity, most betas equal 0). We considered TFBS clusters to have the best evidence for association with rank-responsive genes if they (i) had a non-zero beta in more than half of the 1000 elastic net iterations, and (ii) a Fisher's Exact Test B-H corrected p-value < 0.05.

Using these criteria, we identified multiple enriched TFBS that predicted whether a gene was significantly positively or negatively associated with dominance rank with 62.5% (range: 60.8-64.5%) and 58.8% (range: 52.5%-60.5%) accuracy in the natural killer cells and helper T cells, respectively (fig. S10). In the NK cells, many of the transcription factors that were most predictive of a gene being more highly expressed in low status individuals are linked to inflammatory processes, including NFkB (Fisher's Test OR = 4.01, p= 3.3×10^{-5}), GLI-3, and (Fisher's Test OR = 2.44, p= 1.3×10^{-5}), and SRF (Fisher's Test OR = 8.91, p= 5.3×10^{-4}; fig. S6 and Table S7)

Text S11. Deconvolution of LPS challenge data and 'reconvolution" of PBMC gene expression levels from cell type-specific data

To test that the cell types responsible for gene expression responses to LPS stimulation conformed with expectations (i.e., were largely driven by monocytes and granulocytes), we performed a deconvolution analysis of the gene expression data collected in the LPS stimulation experiment. To do so, we drew on the cell type composition data measured via flow cytometry for each sample (fig S9), and grouped them into six main cell populations: (1) $CD14^+$ monocytes; (2) $CD16^+$ NKs; (3) B cell; (4) helper T cells; (5) cytotoxic T cells; and (6) granulocytes. We then applied the least-squares, partial deconvolution approach implemented in R package csSAM (*50*). We ran csSAM separately for the gene expression data from control versus LPS+ samples (in both cases after controlling for batch effects). We then tested for differential gene expression between conditions in each of the eight deconvolved datasets using a Welch-Satterthwaite t-test with a Benjamini-Hochberg FDR correction for multiple testing. Finally, we performed Fisher's Exact Tests to test whether LPS-responsive genes were enriched within the set of LPS-responsive genes identified for each de-convolved gene expression data set (fig S7: cell lines with median proportion >0.05 in whole blood are shown). Only granulocytes ($\log_2(OR)=1.15$, $p<10^{-63}$) and monocytes ($\log_2(OR)=0.46$, $p=5.7 \times 10^{-12}$) were significantly enriched for the LPS responsive genes identified in the full sample (fig. S7).

To test whether the control samples in the LPS challenge experiment were consistent with our results for the purified, cell type-specific gene expression data, we also performed an *in silico* "reconvolution" analysis for each individual. This analysis was based on the gene expression data for each purified cell type and the flow cytometry-based estimates of cell type proportions. Specifically, we calculated the reconvoluted expression level for each gene-individual combination as the mean of the expression level

of the gene across purified cell types (helper T, cytotoxic T, monocyte, NK, and B), weighted by the proportion of each cell type in the PBMC pool for a given sample. For each female/phase combination:

$$Y = pX \quad (6)$$

where p is a 1 by 5 vector of the proportion of each of the 5 cell types; X is a 5 by n gene matrix of the gene expression values, for the n genes measured (in counts per million, CPM) in the cell type-specific gene expression data (one column for each cell type); and Y is a 1 by n gene vector of the "reconvoluted" gene expression values for the PBMC pool. We calculated PBMC gene expression level estimates for all 45 females for each phase separately and modeled the reconvoluted expression values as a function of rank and age using the same mixed effects models as in the cell type-specific analysis. Note that our reconvoluted estimates differ somewhat from the control samples in the LPS challenge experiment because we included only PBMC cell types in these estimates, whereas in the LPS challenge data, we profiled gene expression levels for all white blood cells, including polymorphonuclear cells.

We found that the effect of rank on the reconstituted gene expression was significantly positively correlated with the effect of rank on gene expression in white blood cells in the control condition of the LPS experiment (Pearson's r = 0.35, $p<2.50\text{x}10^{-248}$). This relationship was stronger when we only considered rank-responsive genes detected in the control condition of the LPS experiment at an FDR < 1% (Pearson's r = 0.59, $p=7.14\text{x}10^{-165}$).

Text S12. Effects of dominance rank on the magnitude of the response to LPS stimulation

As described in equations (3) and (4), rank effects on gene expression in the LPS challenge data are estimated separately for each condition (control versus LPS+). This formulation allowed us to estimate the LPS effect on gene expression for high-ranking and low-ranking females separately as:

$$\beta_{LPS}(r) = \delta + \beta_{rxc} r \quad (7)$$

where $\beta_{LPS}(r)$ is the effect of the LPS treatment for a given Elo rating r, δ is the LPS treatment effect estimated in equation (4), and β_{rxc} is the interaction effect between rank and condition, also estimated in equation (4). In the main text, we calculated these values, across all LPS-responsive genes, for r equal to the mean Elo rating for the lowest ranking females across groups and for r equal to the mean Elo rating for the highest ranking females across groups. This procedure produced a distribution of LPS treatment effects for low ranking and high ranking females, showing that low-ranking females exhibit exaggerated responses to LPS treatment overall (Fig. 4C, Mann-Whitney test $p<10^{-20}$).

Text S13. Rank effects in MyD88 vs TRIF induced genes.

To investigate rank-dependent polarization of the TLR4 signaling pathways through TRIF versus MyD88-dependent arms, we used the results from (*21*), which identified sets of mouse genes for which a normal response to antigen induction was

either MyD88 or TRIF-dependent. These gene lists were obtained by comparing the gene expression response of macrophages from wild-type mice to MyD88 or TRIF knock-out mice, using six purified TLR agonists. Ramsey et al reported 334 genes that were up-regulated upon stimulation via the MyD88-dependent pathway and 274 by that were up-regulated via the TRIF-dependent pathway, among which 238 and 168 rhesus macaque orthologues were included in our LPS challenge data set, respectively.

Using these annotations, we performed three analyses. First, we tested for a difference in the distributions of dominance rank effect sizes between MyD88-induced versus TRIF-induced genes (conditional on being rank-responsive in the LPS+ condition), using a Mann-Whitney test (Fig. 4B). Second, we tested whether MyD88-induced or TRIF- induced genes were enriched among Category I (upregulated after LPS stimulation and more highly expressed in low-ranking females) or Category II (upregulated after LPS stimulation and more highly expressed in high-ranking females) genes, using a Fisher's Exact Test (Fig. 4C). Third, we asked whether females exhibited systematically different gene expression levels, depending on their social status, in the MyD88 versus TRIF-induced gene sets. For this last analysis, we extracted, for each female, the median gene expression level for all genes in the LPS challenge data set that are up-regulated upon stimulation via the MyD88-dependent pathway (based on Ramsay et al), after controlling for batch effects and other covariates (age, cell composition). We then performed linear regression of Elo rating versus median MyD88-dependent gene expression levels across females (Fig. 4D). We repeated the same analysis for genes up-regulated via the TRIF-dependent pathway.

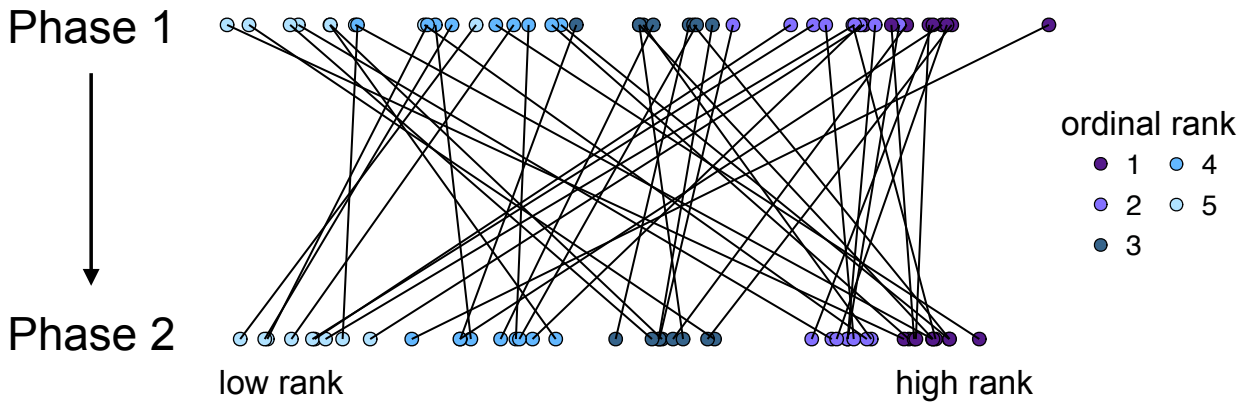

Figure S1. Correlation between female dominance ranks across phases. Solid lines connect the same female in the two phases of the study, with dots colored by ordinal rank in each phase. There was no significant correlation between female ranks across Phases (Pearson's r=0.06, p = 0.68 between phases).

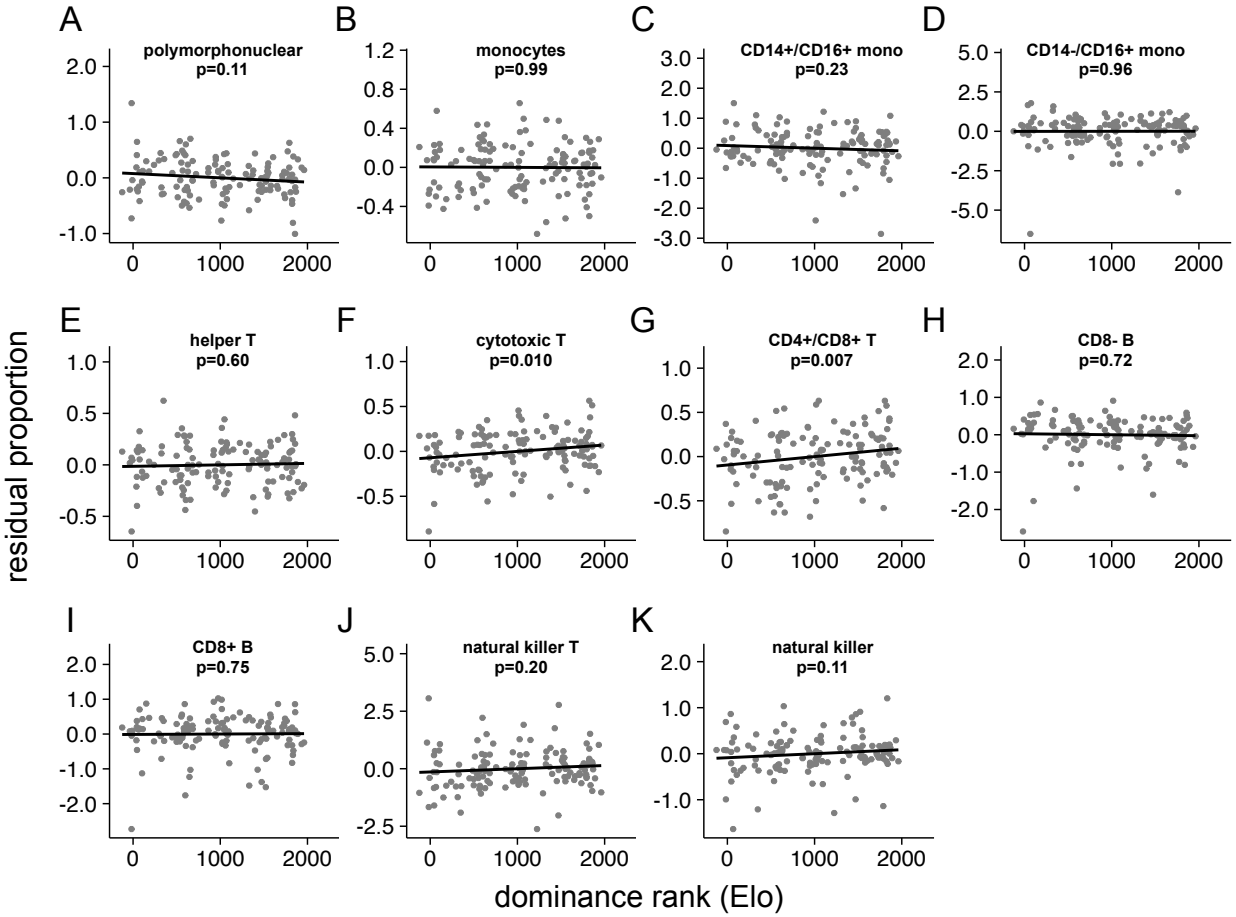

Figure S2. Association between dominance rank and proportions of 11 white blood cell populations. Each plot shows the relationship between Elo rating and the logit-transformed residual proportions of 11 blood cell populations after controlling for subject age and measurement round (proportion data was collected once in Phase 1 and twice in Phase 2, for three repeated measures per individual). **(A)** polymorphonuclear ($CD14^{dim}$/SSC-A>100K/FSC-A>100K), **(B)** classical monocytes ($CD14^{+}/CD16^{-}$), **(C)** $CD14^{+}$ intermediate monocytes ($CD14^{+}/CD16^{+}$), **(D)** $CD14^{-}$ activated non-classical monocytes ($CD14^{-}/CD16^{+}$), **(E)** helper T cells ($CD3^{+}/CD4^{+}$), **(F)** cytotoxic T cells ($CD3^{+}/CD8^{+}$), **(G)** double positive T cells ($CD3^{+}/CD4^{+}/CD8^{+}$), **(H)** $CD8^{-}$ B cells ($CD3^{-}/CD20^{+}/CD8^{-}$), **(I)** $CD8^{+}$ B cells ($CD3^{-}/CD20^{+}/CD8^{+}$), **(J)** natural killer T lymphocytes ($CD3^{+}/CD16^{+}$), and **(K)** natural killer cells ($CD3^{-}/CD16^{+}$)

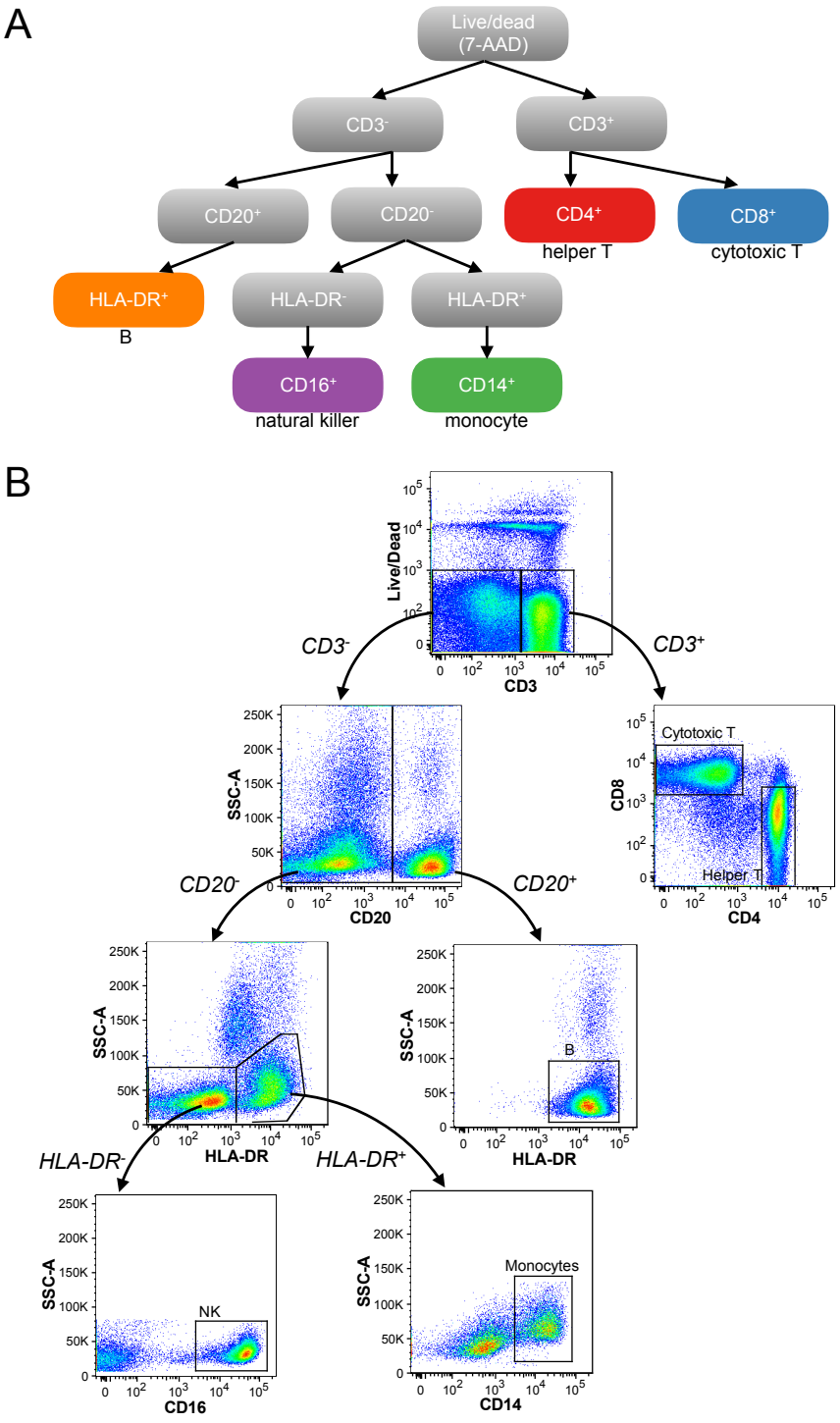

Figure S3 Cell sorting strategy. (A) Schematic of the cell sorting strategy for purifying five populations of PBMCs prior to gene expression analysis. We first removed all dead cells and then sorted the cells using 7 cell surface markers (detailed in the Materials and Methods and Table S2). **(B)** Example of the sorting strategy for one sample visualized using FlowJo (FlowJo, LLC, Ashland, OR).

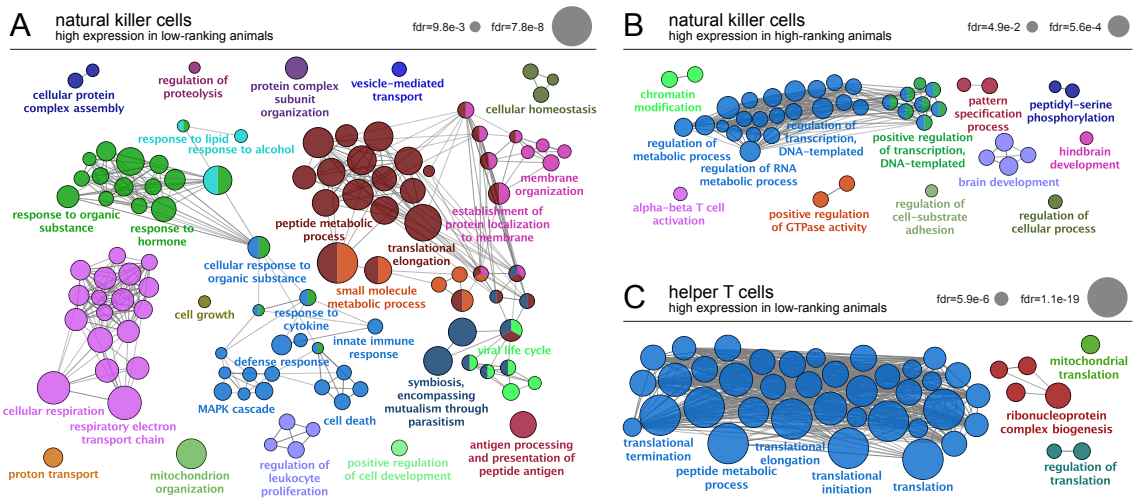

Figure S4 Gene Ontology enrichment in the cell type-specific analysis. (A) Categories enriched in genes more highly expressed in low-ranking animals in NK cells. **(B)** Categories enriched in genes more highly expressed in high-ranking animals in NK cells. **(C)** Categories enriched in genes more highly expressed in low ranking animals in helper T cells. We detected only two significant terms for genes that were more highly expressed in high-ranking individuals in helper T cells, so have not plotted them here (regulation of nervous system development, FDR-corrected $p=7.0 \times 10^{-4}$; regulation of vesicle-mediated transport, $p=3.2 \times 10^{-3}$).

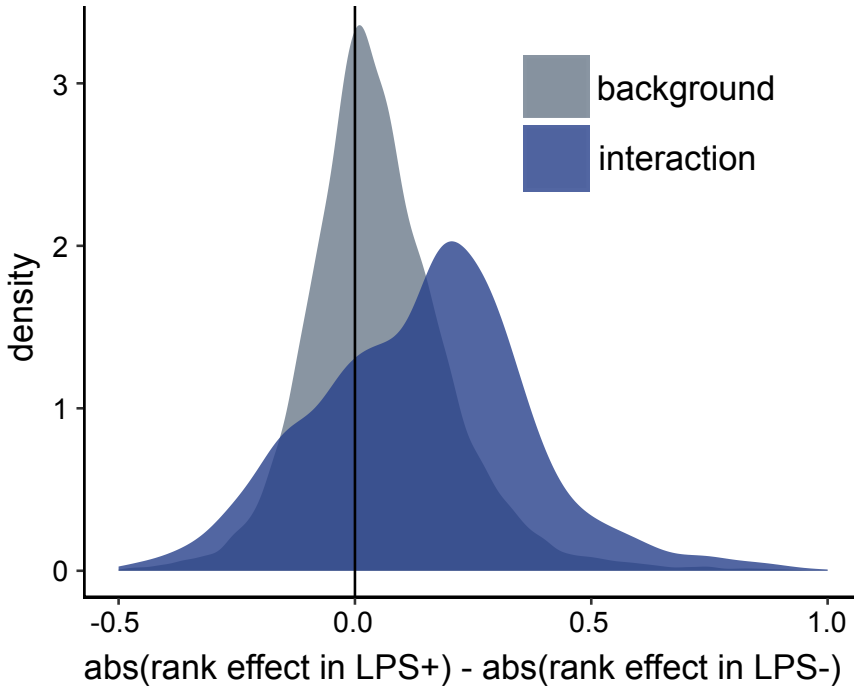

Figure S5. Distribution of the absolute difference in magnitude of rank effects between LPS-stimulated and untreated samples. Gray: genome wide distribution. Blue: genes where we detected a significant interaction effect between condition (control versus LPS+) and dominance rank (FDR < 0.05). Interaction effects are biased towards systematically larger effects of rank after LPS stimulation (Mann-Whitney test, $p=4.2x10^{-86}$).

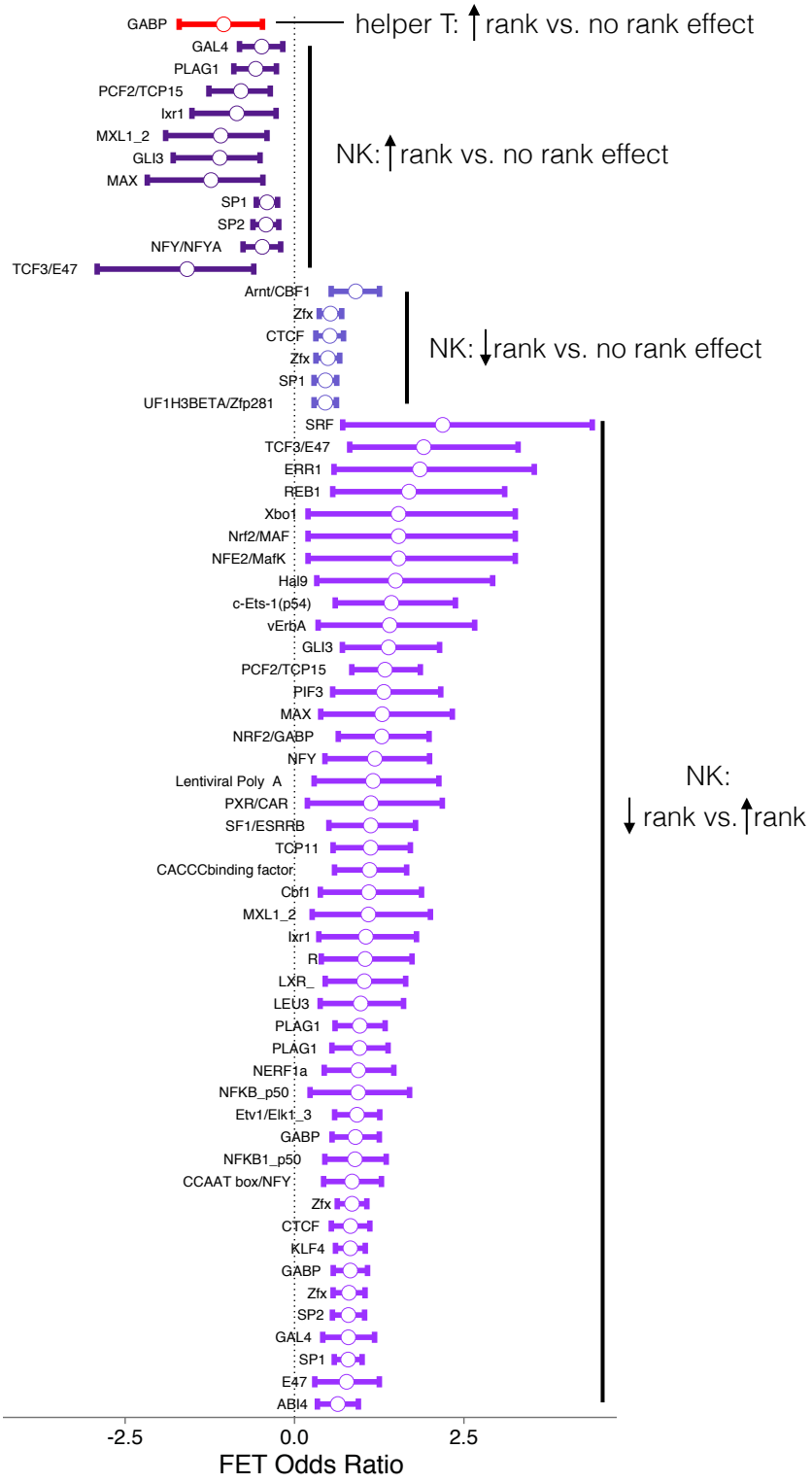

Figure S6. Transcription factors with putative binding sites enriched near rank-responsive genes. FET odds ratios for TFBS clusters in which (i) the FET results were significant at a B-H FDR < 5%, and (ii) presence of the TFBS in open chromatin within 5 kb of gene transcription start sites contributed to predicting whether genes were rank-responsive or more highly expressed in high versus low-status females (non-zero betas in an elastic net logistic regression, for at least 50% of test sets). Points and whiskers show FET OR ± 95% CI.

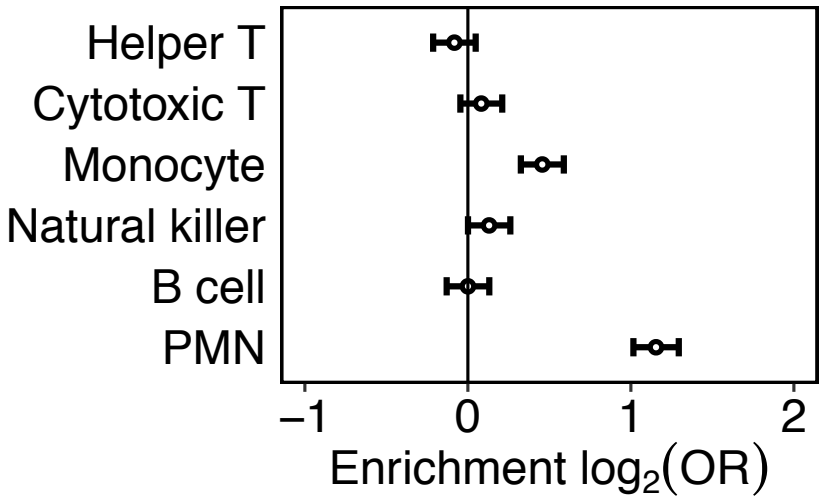

Figure S7. Enrichment of LPS-responsive genes in deconvoluted cell type-specific gene expression data. We tested for enrichment of LPS-responsive genes identified in the deconvoluted data among the set of LPS-responsive genes identified in the non-deconvoluted empirical data. In both cases, LPS-responsive genes were defined as genes differentially expressed between the control and LPS+ conditions at an FDR<0.01. As expected, granulocytes ($\log_2$(OR)=1.15, p=$8.7\text{x}10^{-64}$) and monocytes ($\log_2$(OR)=0.46, p=$5.7\text{x}10^{-12}$) were most important in mediating this response.

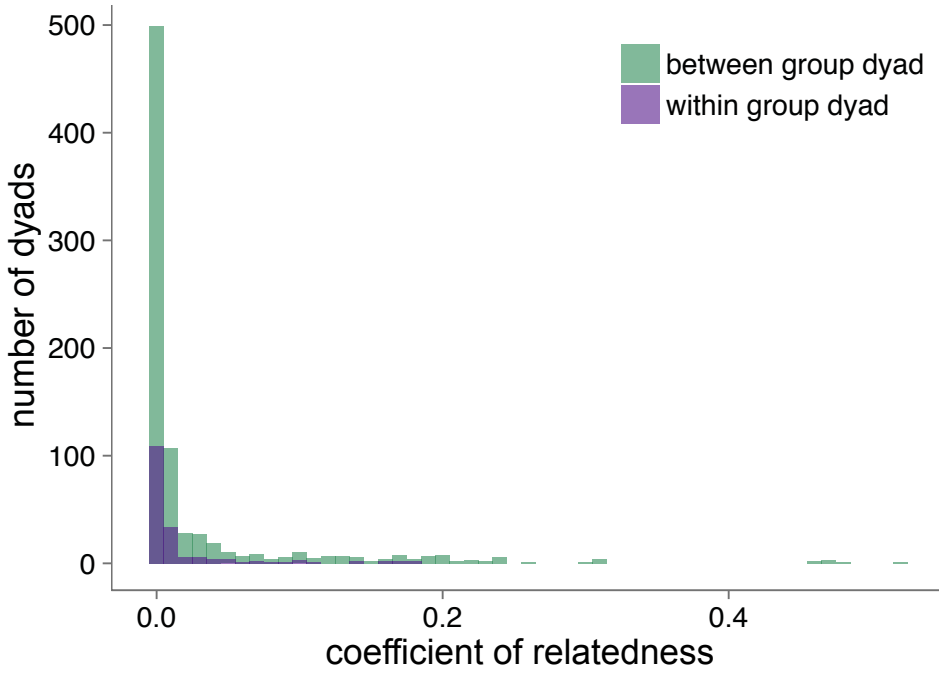

Figure S8. Kinship among study subjects. No groups contained first-degree relatives. The x-axis depicts the coefficient of relatedness (*r*), which ranges from 0-1 (r = 0.5 for full siblings and parent-offspring dyads).

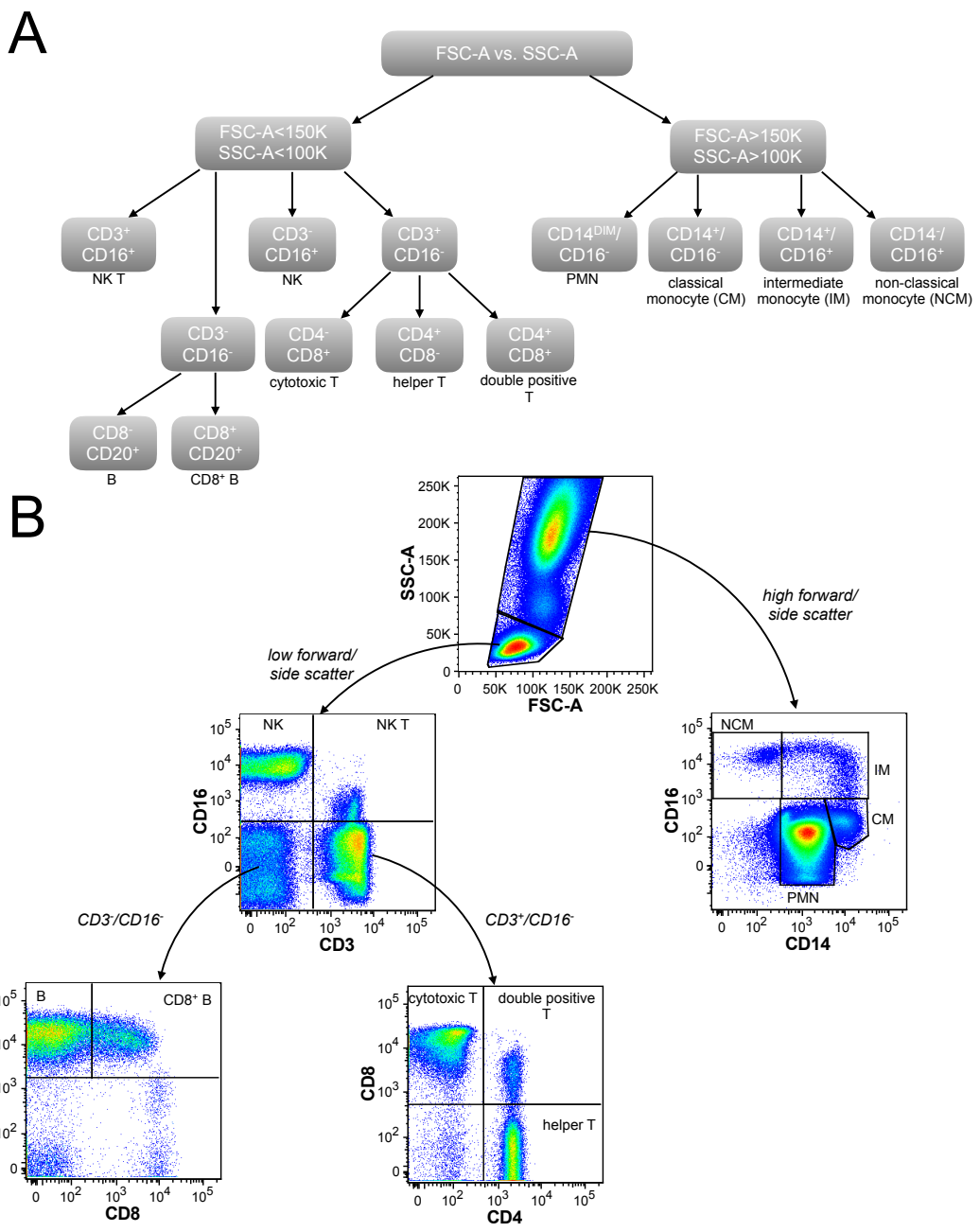

Figure S9. LPS challenge experiment cell phenotyping strategy. (A) Schematic of the gating strategy for phenotyping the cells in the LPS challenge experiment. We first gated cells based on forward and side scatter, and then classified them into 11 cell populations using a combination of cell surface markers (Table S2). **(B)** Example of the phenotyping strategy for one sample visualized using FlowJo (FlowJo, LLC, Ashland, OR).

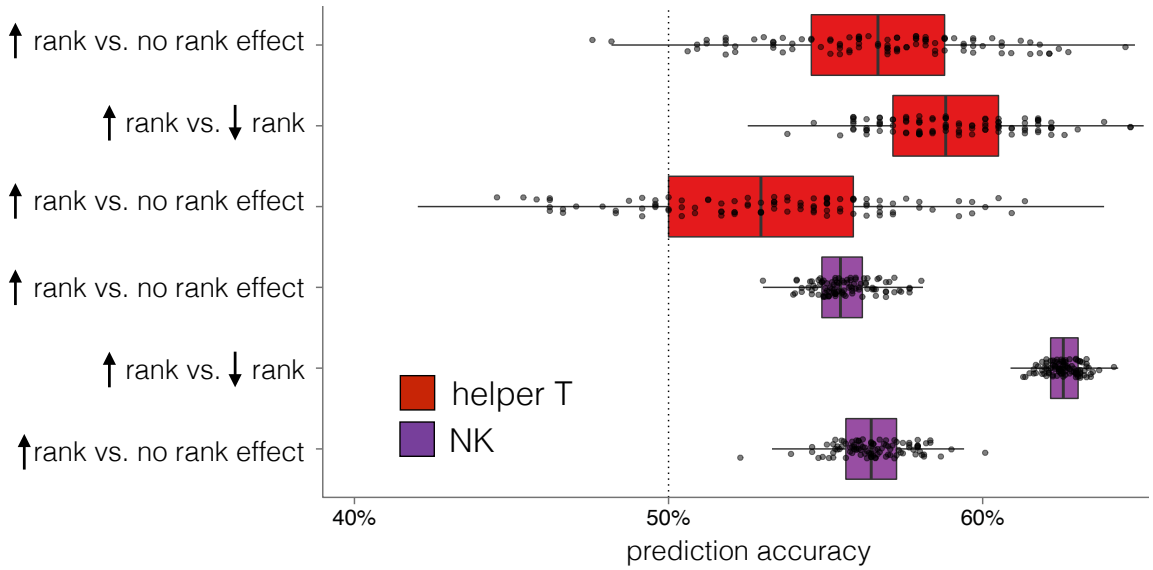

Figure S10. TFBS-based prediction of rank-responsive genes. Using putative TF binding sites (TFBS) in regions of open chromatin 5 kb upstream of the TSS of a gene, we could predict whether that gene was significantly positively (↑ rank) or negatively (↓ rank) associated with dominance rank with 62.5% (range: 60.8-64.5%) and 58.8% (range: 52.5%-60.5%) accuracy in natural killer cells and helper T cells, respectively.